



April 3, 2025

VIA ECF

Honorable Ona T. Wang
United States District Court
500 Pearl Street
New York, NY 10007

Re: *Authors Guild v. OpenAI Inc.*, No. 23-cv-8292, and *Alter v. Open AI Inc.*, 23-cv-10211 (S.D.N.Y.): Response to Brief Regarding Microsoft's Response to Interrogatory No. 8 ([ECF 346](#))

Dear Magistrate Judge Wang:

Plaintiffs' Interrogatory No. 8 requests a list of every individual within Microsoft who has ever had any communication about nearly a dozen discrete items, including the popular website Reddit, the Internet Archive, and others. In response to Microsoft's well-founded relevance, overbreadth, and burden objections, Plaintiffs now purport to "narrow" their interrogatory to a list of all individuals who have ever had any communication, regardless of subject matter, regarding "LibGen." But Microsoft's objections to the original interrogatory are not resolved by simply focusing on one subject matter that is just as problematic as the rest.

The Court should reject Plaintiffs' motion for at least three reasons. First, on its face, this interrogatory is overbroad and irrelevant because it demands that Microsoft investigate the communications of every single one of its more than 200,000 employees for anything about LibGen, whether or not any communication is relevant to the issues in these consolidated cases. Second, Microsoft has already provided discovery about any relevant references to LibGen by existing custodians, which should satisfy any reasonable version of its interrogatory. Finally, to the extent Plaintiffs' relevance argument attempts to reach more broadly than OpenAI's LLM training, then this interrogatory would be a backdoor request for reconsideration of the Court's previous orders limiting the scope of discovery in this case.

Plaintiffs' interrogatory seeks information far outside the boundaries of relevance and proportionality set by Rule 26(b)(1). It asks Microsoft to identify *every* one of its over 200,000 employees—including those hired yesterday—who has communicated about LibGen—*ever*, in their lives, across time and space, for *any* reason.¹ Even if Plaintiffs were to limit the number or scope of individuals, the interrogatory is still overbroad. It asks, simply, who has communicated about LibGen. Not whether someone has communicated about LibGen for Microsoft-related purposes. Not whether someone has communicated about LibGen while working for Microsoft. Not even whether someone has communicated about LibGen with respect to something related to

¹ Searching all Microsoft employees' files for such communications would also be an end-run around the parties' agreed process for identifying specific custodians and then searching their files for potentially responsive documents as set out in the Stipulation and Order Re: Discovery of Electronically Stored Information. See ECF 201 at 2-3.

Honorable Ona T. Wang

- 2 -

April 3, 2024

AI. And certainly not whether someone has communicated about LibGen with respect to the training of OpenAI's models, which is the conduct alleged in these cases.

Plaintiffs now suggest in their motion that perhaps they simply want the names of individuals who have communicated about LibGen as discovered through Microsoft's investigation to date.² But Plaintiffs already have that information with respect to OpenAI's LLM training. Microsoft has searched communications of the custodians agreed on by the parties for LibGen-related terms, including the search proposed by Plaintiffs: LibGen OR librarygenesis OR "library genesis" OR "lib gen." Microsoft has produced and will produce responsive documents that hit on this term, as relevant to the training of OpenAI's at-issue models. Plaintiffs can easily discern the senders and recipients of any such communications from a simple review of the documents that have been produced in this case.

Finally, to the extent Plaintiffs' relevance argument attempts to reach more broadly than OpenAI's LLM training, the interrogatory is a backdoor attempt to get around the Court's earlier orders limiting discovery. *See* ECF 293 at 3 (denying Plaintiffs' motion filed at ECF 269 (sealed), ECF 271 (redacted), which moved to compel, *inter alia*, documents related to Microsoft's LLMs).

For all these reasons, Class Plaintiffs' request to compel Microsoft to respond to Interrogatory No. 8 should be denied.

Respectfully submitted,

Respectfully submitted,

/s/ Annette L. Hurst/s/ Jared B. Briant

Annette L. Hurst

Jared B. Briant

Counsel for Defendant Microsoft Corporation

² Plaintiffs suggest they "emphasized they would accept a list of the individuals that Microsoft's reasonable investigation has revealed were communicating about LibGen." ECF 346 at 1. But Plaintiffs did not offer this as a proposal at the meet and confer—at best, they noted it as a limitation they could potentially consider—and it was absent from their follow-up email. *See* ECF 346-3, Ex. C. at 7. Plaintiffs do not seem to propose this "narrowed" version now: it is only mentioned in the Background section of Plaintiffs' brief, and not in the request to the Court as articulated in the first or last sentences. *See* ECF 346. In any event, the interrogatory still is not limited to the allegations or issues in the case, and this proposal introduces new problems because it would have Microsoft reveal attorney-client privileged communications and work product by revealing specifics of what was discussed by litigation counsel with particular Microsoft witnesses.